Cell, Vol. 110, 339–348, August 9, 2002, Copyright ©2002 by Cell Press

Inappropriate Gene Activation in FSHD: A Repressor Complex Binds a Chromosomal Repeat Deleted in Dystrophic Muscle

Davide Gabellini,[1] Michael R. Green,[1] and Rossella Tupler[1,2,3]
[1]Howard Hughes Medical Institute
Program in Gene Function and Expression
Program in Molecular Medicine
University of Massachusetts Medical School
Worcester, Massachusetts 01605
[2]Biologia Generale e Genetica Medica
Universita' degli Studi di Pavia
27100 Pavia
Italy

Summary

Facioscapulohumeral muscular dystrophy (FSHD), a common myopathy, is an autosomal dominant disease of unknown molecular mechanism. Almost all FSHD patients carry deletions of an integral number of tandem 3.3 kilobase repeats, termed D4Z4, located on chromosome 4q35. Here, we find that in FSHD muscle, 4q35 genes located upstream of D4Z4 are inappropriately overexpressed. We show that an element within D4Z4 specifically binds a multiprotein complex consisting of YY1, a known transcriptional repressor, HMGB2, an architectural protein, and nucleolin. We demonstrate that this multiprotein complex binds D4Z4 in vitro and in vivo and mediates transcriptional repression of 4q35 genes. Based upon these results, we propose that deletion of D4Z4 leads to the inappropriate transcriptional derepression of 4q35 genes resulting in disease.

Introduction

Facioscapulohumeral muscular dystrophy (FSHD) is the third most common inherited myopathy, with an incidence of 1:20,000 (Padberg, 1982). The disease is characterized by progressive weakness and atrophy of the facial and shoulder girdle muscles, which subsequently spreads to the abdominal and pelvic girdle muscles with highly variable expression. The genetic defect follows autosomal dominant inheritance, although new mutations account for approximately 10% of recognized cases (Lunt, 1998).

The FSHD locus was mapped to the subtelomeric region of the long arm of chromosome 4, 4q35, by genetic linkage analysis (Sarfarazi et al., 1992). Almost all FSHD patients carry rearrangements occurring in a 3.3 kilobase (kb) tandemly repeated sequence (D4Z4) located at the 4q subtelomeric region (Wijmenga et al., 1992). These rearrangements result in EcoRI alleles shorter than 35 kb, providing the molecular marker for FSHD diagnosis (Lunt, 1998).

D4Z4 is highly polymorphic with a variable number tandem repeat (VNTR) structure (Hewitt et al., 1994; Winokur et al., 1994). In the general population, its size may vary between 11 and 150 units, whereas FSHD patients carry fewer than 11 repeats (Lunt, 1998). The number of D4Z4 repeats is a critical determinant of the age of onset and clinical severity of FSHD (Goto et al., 1995; Lunt et al., 1995; Zatz et al., 1995; Tawil et al., 1996; Hsu et al., 1997; Ricci et al., 1999). In general, 1–3 D4Z4 repeats is associated with a severe form of disease that presents in childhood, 4–7 repeats with the most common form of FSHD, and 8–10 repeats with a milder disease and reduced penetrance.

Two candidate genes, *FRG1* and *TUB4q*, were isolated from the FSHD critical region, but their role in FSHD pathogenesis has failed to be substantiated (van Deutekom et al., 1996; van Geel et al., 2000). Recently, a new gene, *FRG2*, mapping 37 kb proximal to D4Z4, has been discovered (R.R. Frants, personal communication), although its biological function has not been elucidated. In addition, a putative open reading frame (*DUX4*) within D4Z4 has been identified (Gabriels et al., 1999), but there is no evidence that it is expressed.

Some normal individuals carry an abnormal chromosome 4 resulting from an unbalanced translocation between the 4q35 subtelomeric region and the short arm of an acrocentric chromosome (Tupler et al., 1996). The rearranged chromosome 4 lacks the entire D4Z4 repeat and a 200 kb proximal region including *FRG2*, *TUB4q*, and *FRG1*. Thus, haploinsufficiency of the entire 4q subtelomeric region has no phenotypic consequence, whereas deletion of only the D4Z4 repeats is associated with FSHD. Collectively, these results suggest that FSHD results from a gain-of-function mutation and raise the possibility that 4q35 genes proximal to D4Z4 play a role in disease initiation or progression.

These observations strongly suggest that FSHD is not the result of a classical mutation within a protein-coding gene. We have previously reported that in FSHD muscle many genes are transcriptionally misregulated, raising the possibility that a regulatory defect might be the underlying basis of disease (Tupler et al., 1999). Here we investigate the possibility that deletion of D4Z4 repeats initiates transcriptional misregulation.

Results

Inappropriate Overexpression of 4q35 Genes in FSHD Dystrophic Muscle

To investigate whether altered expression of 4q35 genes could be the underlying basis of FSHD, we collected human muscle samples from normal individuals and patients affected by FSHD as well as several other muscular dystrophies including Duchenne muscular dystrophy (DMD), amyotrophic lateral sclerosis (ALS), and limb girdle muscular dystrophy type 2 (LGMD). Expression of three 4q35 genes, FSHD region gene 1 (*FRG1*), FSHD region gene 2 (*FRG2*), and the adenine nucleotide translocator-1 gene (*ANT1*), was analyzed by RT-PCR.

Figure 1A shows that expression of *FRG2*, the gene most proximal to D4Z4, was undetectable in normal muscle but was present at a significant level in all three FSHD muscle samples. Expression of the two other 4q35

[3]Correspondence: rossella.tupler@umassmed.edu

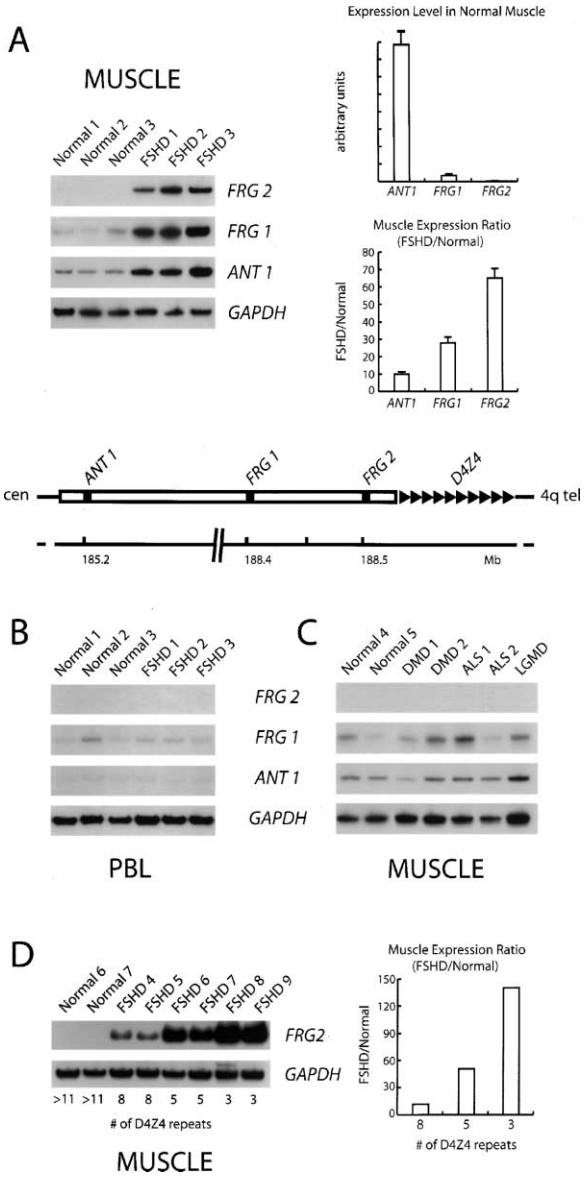

Figure 1. Inappropriate Overexpression of 4q35 Genes in FSHD Dystrophic Muscle

(A) *FRG2*, *FRG1*, and *ANT1* expression analysis was performed on RNA extracted from muscles derived from three normal individuals (Normal 1–3) and three FSHD patients (FSHD 1–3) by radioactive RT-PCR using specific primers. Results are presented for three of three total FSHD patients analyzed. *GAPDH* expression was also monitored and used for sample normalization. Signal intensity was quantitated by phosphoimager analysis; the expression level of each gene in normal muscle (top right) and the ratio of expression levels in FSHD and normal muscle (bottom right) are shown. (Bottom) Schematic representation of 4q35 showing physical distances between the genes in megabases (Mb), which were established based on the chromosome map published at http://www.ncbi.nlm.nih.gov, and the localization of *FRG1*, *FRG2* and D4Z4 within the AF146191-U85056 contig.

(B) *FRG2*, *FRG1*, *ANT1*, and *GAPDH* expression analysis was performed on RNA extracted from peripheral blood lymphocytes (PBL) of the same normal and FSHD patients as in (A) by radioactive RT-PCR.

(C) Analysis of *FRG2*, *FRG1*, *ANT1*, and *GAPDH* expression in muscle from two other normal individuals (Normal 4 and 5) and patients affected by Duchenne Muscular Dystrophy (DMD), Amyotrophic Lateral Sclerosis (ALS), and Limb Girdle Muscular Dystrophy type 2 (LGMD) patients.

genes, *FRG1* and *ANT1*, was detected in all muscle samples but overexpressed only in FSHD muscle. Quantitation by phosphoimager analysis revealed, first, that in normal muscle the expression level of 4q35 genes increased with distance from D4Z4 and, second, that in FSHD muscle the level of overexpression varied inversely with distance from D4Z4. Glyceraldehyde Phosphate Dehydrogenase (*GAPDH*), which is not at 4q35, was expressed equivalently in all muscle samples. Significantly, Figure 1B shows that in lymphocytes from FSHD patients, expression of *FRG2*, *FRG1*, and *ANT1* was equivalent to that observed in normal tissue, indicating that overexpression of 4q35 genes in FSHD is muscle specific. Expression of 4q35 genes was also normal in DMD, ALS, and LGMD muscle samples (Figure 1C), indicating that misregulated 4q35 gene expression is not a general characteristic of muscular dystrophies.

As described above, the number of D4Z4 repeats is a critical determinant of disease severity (Ricci et al., 1999). The results of Figures 1A–1C raised the possibility that inappropriate overexpression of 4q35 genes located upstream of D4Z4 was the underlying basis of FSHD. A prediction of this model is that 4q35 gene overexpression would be increased when the number of D4Z4 repeats is decreased. To test this prediction, we analyzed expression of *FRG2* in muscle biopsies from six additional FSHD patients carrying varying numbers of D4Z4 repeats. The results of Figure 1D clearly demonstrate that the level of *FRG2* overexpression is indeed inversely related to the number of D4Z4 repeats. On the basis of Figures 1A–1D, we conclude that 4q35 gene mRNAs and presumably proteins are overexpressed in FSHD muscle in a manner inversely related to distance from D4Z4 and the number of D4Z4 repeats.

Detection of a Nuclear Activity that Binds to D4Z4

The most likely mechanism by which D4Z4 could suppress 4q35 gene transcription is through interaction with a cellular factor (or factors) that silences the expression of nearby genes. We tested this possibility by analyzing the interaction between D4Z4 and nuclear proteins using an electrophoretic mobility shift assay (EMSA). Probes were generated by ^{32}P-end-labeling eight restriction enzyme-digested fragments that spanned the entire D4Z4 sequence. Figure 2A shows that one DNA fragment, termed D4Z4-243, supported formation of a specific complex. This complex was abolished by addition of increasing amounts of D4Z4-243 DNA but not by nonspecific DNA (Figure 2B). A comparable D4Z4-243 binding activity was also detected in nuclear extracts prepared from several human and mouse myogenic cell lines (see below and data not shown).

Deoxyribonuclease I (DNase I) footprinting was used to map the binding site within D4Z4-243. Figure 2C shows that following incubation with nuclear extracts

(D) Analysis of *FRG2* and *GAPDH* expression in muscle from two additional normal (Normal 6 and 7) and six additional FSHD patients (FSHD 4–9) carrying the indicated number of D4Z4 repeats. Signal intensity was quantitated by phosphoimager analysis and normalized based upon to *GAPDH* signal; the ratios of expression levels between FSHD muscles and normal muscle (right) are shown.

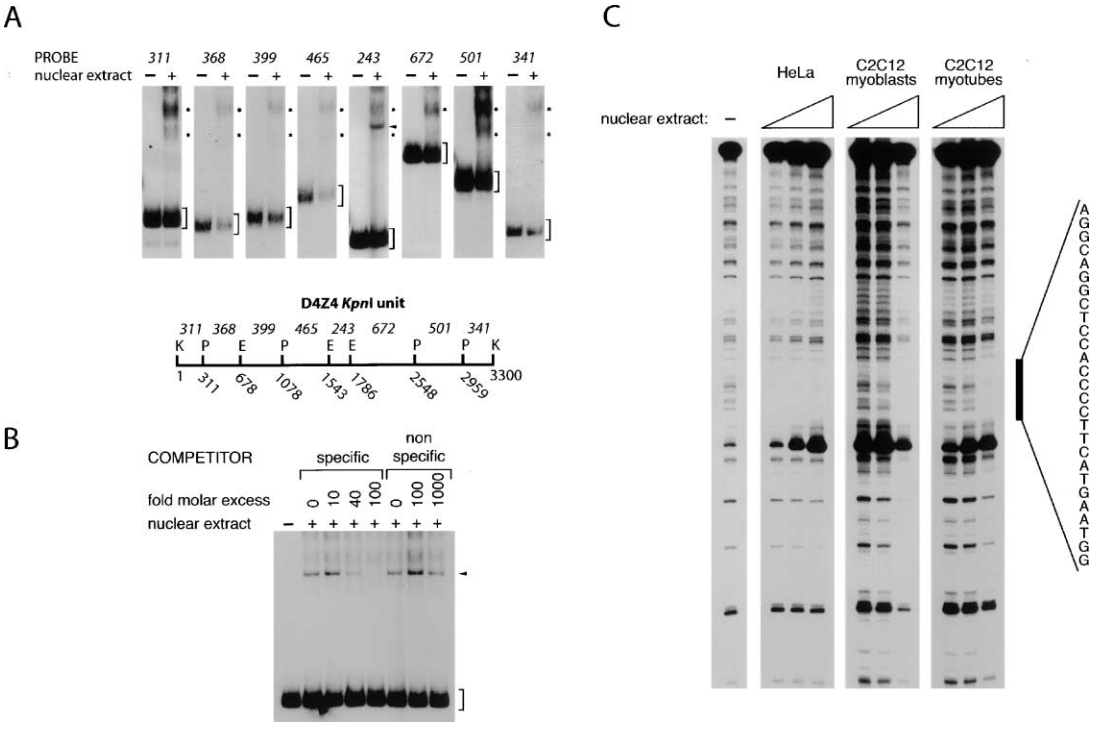

Figure 2. Detection of a Nuclear Activity that Binds to D4Z4

(A) Electrophoretic mobility shift assay (EMSA). Each fragment was end-labeled, incubated with (+) or without (−) HeLa nuclear extract for 30 min, and fractionated on nondenaturing polyacrylamide gels. The specific shift with probe 243 is indicated by the arrowhead; the nonspecific shifts are indicated by asterisks; the position of the free probe is indicated by the bracket. (Bottom) Schematic representation of the 3.3 kb D4Z4 KpnI unit, showing the positions of the restriction enzyme sites KpnI (K), PstI (P), and EagI (E), as well as the fragment sizes generated by enzyme digestion.

(B) The D4Z4-243 probe was incubated with (+) or without (−) nuclear extract, in the presence of molar excess of cold specific or nonspecific competitor as indicated. The mobility shift is indicated by the arrowhead, and the position of the free probe is indicated by the bracket.

(C) In vitro DNase I footprinting. D4Z4-243 was end-labeled and incubated with or without DNase I for 1 min in the presence or absence of increasing quantities of either HeLa, C2C12 myoblast, or C2C12 myotube nuclear extract. The black bar represents the protected region, and the corresponding 27 bp sequence is indicated. A G+A sequencing ladder (not shown) run in parallel with the footprinting reactions was used to orient the sequence.

from HeLa cells, C2C12 myoblasts, or C2C12 myotubes, a 27 basepair (bp) sequence (CCATTCATGAAGGGGTG GAGCCTGCCT) within D4Z4-243 was protected from DNase I digestion. We will refer to this 27 bp sequence as the D4Z4 binding element (DBE).

The DBE Is a Transcriptional Repression Element

One explanation for the results presented above is that the DBE is a transcriptional repression element that, when deleted in FSHD, leads to overexpression of 4q35 genes. To test this hypothesis, we analyzed the ability of the DBE to suppress transcription of a reporter gene expressed from the strong cytomegalovirus (CMV) promoter. Analogous assays have been used previously to identify positive and negative transcriptional regulatory elements (see, for example, Chung et al., 1997).

We constructed a series of plasmids in which increasing numbers of DBEs were inserted between the CMV promoter and the neomycin-resistance gene. The plasmids were stably transfected into HeLa cells, and the number of G418-resistant colonies was quantitated. As shown in Figure 3, the number of inserted DBEs inversely correlated with the number of G418-resistant colonies: in the presence of one, two, or six DBEs, the number of G418-resistant colonies decreased by 25%, 45%, and 86%, respectively. Significantly, insertion of a nonrelated spacer sequence had no effect, indicating that the decrease in colony number was not simply due to increased distance between the CMV promoter and the reporter gene. On the basis of these results, we conclude that the DBE within D4Z4 is a transcriptional repression element.

Identification of a Multiprotein Complex that Binds to the DBE

To identify and characterize proteins that bind to the DBE, we biochemically purified the D4Z4 recognition complex (DRC) using sequential chromatography of HeLa nuclear extracts on P11 Phosphocellulose, DEAE-Sepharose, and DBE-Agarose (see Figure 4A, schematic). The purified DRC was fractionated on an SDS-PAGE gel, and the polypeptides were detected by silver staining, excised, and microsequenced by mass spectrometry. Figure 4A shows that four bands were present in the active fraction: YY1, a multifunctional repressor/activator (reviewed in Thomas and Seto, 1999); HMGB2, a member of a family of nonhistone chromatin-associated proteins (reviewed in Thomas and Travers, 2001);

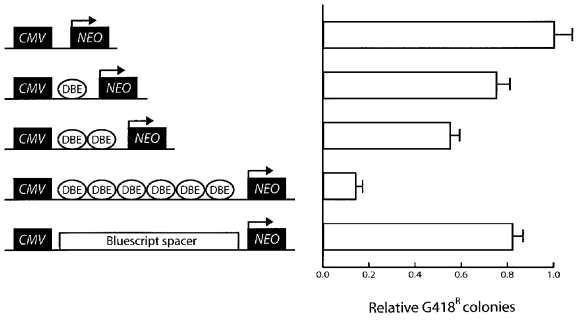

Figure 3. The DBE Is a Transcriptional Repression Element

(Left) Schematic representation of the reporter constructs containing the CMV promoter driving expression of the neomycin-resistance reporter gene (*NEO*). One, two, or six copies of the 27 bp minimal D4Z4 binding element (DBE) were inserted between the CMV promoter and the transcription start site of the *NEO* gene, which is indicated by the arrow. A construct containing a pBluescript spacer equivalent in length to the size of the insert generated by six copies of the DBE was also included as a control.

(Right) HeLa cells were stably transfected and grown in medium containing G418. G418-resistant colonies were counted after 2 weeks; the number of colonies obtained with each construct is shown relative to the number of colonies obtained for the construct lacking D4Z4 binding sites.

nucleolin; and keratin. Keratin is a common contaminant of mass spectrometric analysis (Eng et al., 1994) and was not further analyzed. Immunoblot analysis confirmed the identity of YY1, HMGB2, and nucleolin (Figure 4A).

To verify that this multiprotein complex was in fact the nuclear DBE binding activity, mobility-shift experiments were performed using antibodies directed against YY1, HMGB2, and nucleolin. Figure 4B shows that antibodies to nucleolin and HMGB2 resulted in formation of a lower mobility complex ("supershift"), whereas addition of a control antibody had no effect. An antibody to YY1 interfered with formation of the DBE mobility shift, suggesting that association of YY1 with either a component of the DRC or the DBE itself had been disrupted.

Recognition of the DBE by YY1

We noticed that the DBE contains a putative YY1 recognition sequence (CCATN; Yant et al., 1995), suggesting that YY1 directly interacts with the DBE. To test this possibility, we carried out mobility-shift experiments using recombinant YY1. Figure 4C shows that recombinant GST-YY1, but not GST alone, bound the DBE probe. A DBE probe containing a mutation within the YY1 core recognition sequence was not bound by GST-YY1. Moreover, this mutant probe failed to support a mobility shift with HeLa nuclear extract (Figure 4C, right), suggesting that elimination of YY1 binding also abolished association of the entire DRC to the DBE. We conclude that YY1 has an intrinsic DBE binding activity and that YY1 is the most likely DRC component to directly contact the DBE.

The DRC Binds D4Z4 In Vivo

To determine whether the DRC also binds D4Z4 sequences in vivo, we performed chromatin immunoprecipitation (ChIP) experiments. HeLa cells were first treated with formaldehyde to induce crosslinks between the DBE and bound proteins. DNA-protein complexes were then immunopurified using antibodies directed against various DRC components, and the presence of specific DNA fragments in the immunoprecipitate was quantified by PCR. We used three sets of primer pairs that were specific to either D4Z4, p13E-11, or *FRG1*. Figure 5A shows that all three DRC components, YY1, HMGB2, and nucleolin, were associated with D4Z4 sequences. In contrast, there was no detectable binding of any DRC component to the control sequences p13E-11 or *FRG1*.

The human genome contains sequences homologous to D4Z4 on several chromosomes in addition to 4 (Wijmenga et al., 1992; Winokur et al., 1994; Lyle et al., 1995). It was therefore possible that the interaction between the DRC and DBE detected in the ChIP assay shown in Figure 5A did not occur at 4q35, but rather at one or more of these other chromosomal loci. To address this possibility, we used the ChIP assay to analyze binding to the DBE in a human/rodent monochromosomal cell hybrid containing a single human chromosome 4. The rodent genome lacks D4Z4 repetitive sequences (Clark et al., 1996), and therefore in this experimental system the only genomic copy of D4Z4 is present at 4q of the human chromosome. Figure 5B shows that in this cell line, all three DRC components were specifically associated with the D4Z4 sequences. On the basis of these results, we conclude that in vivo the DRC is bound to D4Z4 within 4q35.

Reducing the Levels of DRC Components Results in Overexpression of 4q35 Genes

The results presented above suggest that deletion of D4Z4 repeats reduces the number of DBEs and thus the number of DNA bound DRCs present at 4q35. Reducing the number of bound transcriptional repressing complexes is predicted to result in the inappropriate upregulation of 4q35 genes. To verify that the DRC is responsible for the lack of 4q35 gene expression in normal cells, we performed antisense experiments to decrease the intracellular levels of DRC components.

HeLa cells were transfected with morpholino oligonucleotides targeted to human YY1, HMGB2, or nucleolin mRNA. The immunoblot in Figure 6A shows that each antisense oligonucleotide specifically reduced the level of its cognate protein. Figure 6B shows that reducing the levels of YY1, HMGB2, or nucleolin resulted in overexpression of the 4q35 gene *FRG2*. Thus, reducing the levels of DRC components recapitulates the molecular event observed at 4q35 in FSHD muscle.

Discussion

A Model for the Molecular Basis of FSHD

FSHD is a complex disease with a peculiar involvement of muscle groups, highly variable severity, and unpredictable progression. The disease has been causally related to deletion of subtelomeric D4Z4 repeats at 4q35; however, no candidate gene has been isolated. The failure to identify an expressed sequence in this region has prompted an alternative model that proposes the disease is due to the inappropriate expression of nearby genes by a positional effect. According to this

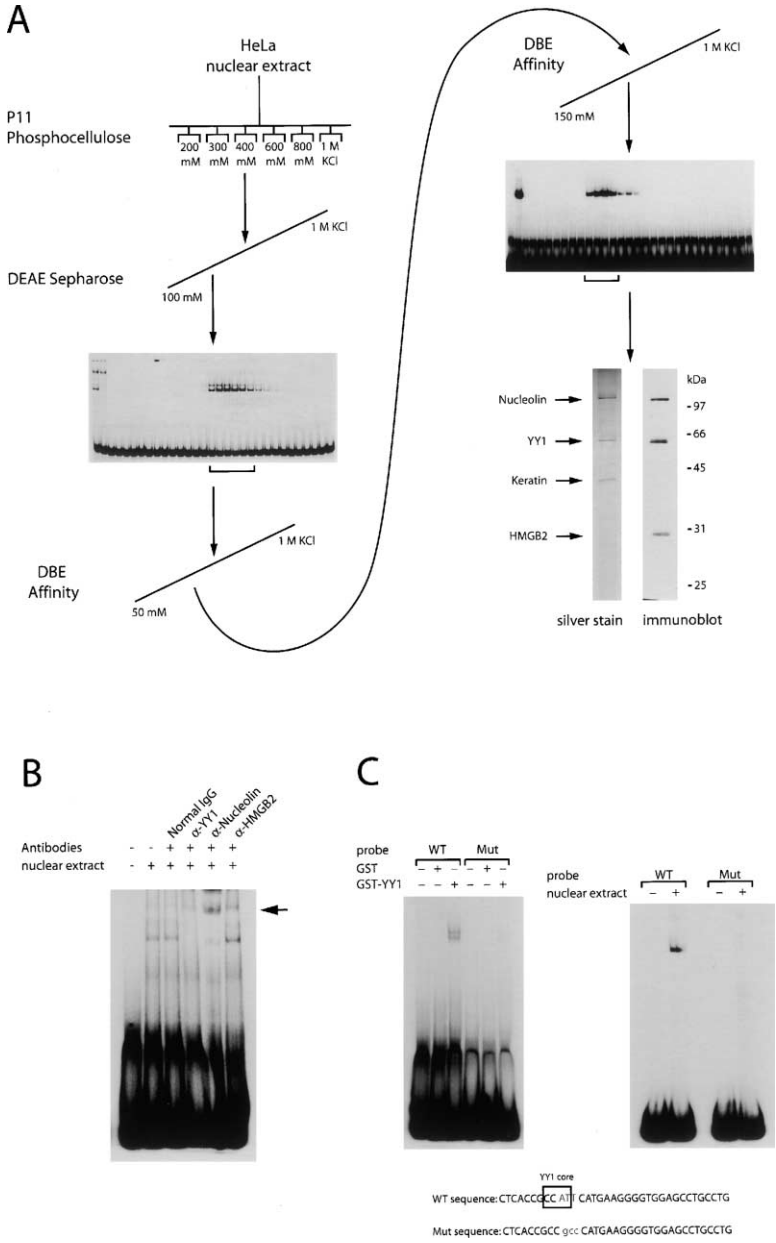

Figure 4. Identification of a Multiprotein Complex that Binds to the DBE

(A) Schematic representation of the chromatographic steps used to purify the D4Z4 binding activity. HeLa nuclear extracts were fractionated first on a P11 Phosphocellulose column, and the active fraction, as monitored by EMSA, was subsequently fractionated on a DEAE-Sepharose column. Active fractions (bracketed lanes) were pooled and loaded onto a DBE affinity resin; the active fractions from this step were pooled and fractionated again on a fresh DBE affinity column. Final active fractions (bracketed lanes) were pooled and separated by SDS-PAGE, and four bands were detected by silver staining (left). Mass spectrometry analysis identified the bands as nucleolin, YY1, keratin, and HMGB2, as shown. The mass spectrometry results were confirmed by immunoblotting (right).

(B) EMSA was performed in the presence of antibodies specific to each DRC component, or in the presence of the control antibody IgG. Anti-YY1 antibodies immunocompeted the mobility shift, while anti-Nucleolin and anti-HMGB2 antibodies generated a supershift (indicated by the arrow).

(C) Recognition of the DBE by YY1 in vitro. EMSA was performed with purified GST-YY1 (left) or nuclear extract (right) using a wild-type (wt) DBE probe or a version mutated in the YY1 recognition sequence, as shown.

model, deletion of D4Z4 places the normally euchromatic genes in 4q35 under the control of telomeric heterochromatin, thereby leading to their transcriptional repression (Fisher and Upadhyaya, 1997).

Based on the results presented in this study, we propose a different model for the molecular basis of FSHD (Figure 7). In normal individuals, the presence of a threshold number of D4Z4 repeats leads to repression of 4q35 genes by virtue of a DNA bound multiprotein complex that actively suppresses gene expression. In FSHD patients, deletion of an integral number of D4Z4 repeats reduces the number of bound repressor complexes and consequently decreases (or abolishes) transcriptional repression of 4q35 genes. As a result, these genes are inappropriately overexpressed, ultimately leading to disease onset and progression. We have shown that the extent of transcriptional derepression is a function of the number of deleted D4Z4 repeats (Figure 1D), which is also a critical determinant of disease severity (Ricci et al., 1999). The 4q35 gene overexpression model of Figure 7 provides a molecular explanation for the autosomal dominant transmission of FSHD. Our finding that 4q35 gene overexpression is muscle specific (Figure 1B) also provides an attractive explanation for the tissue specificity of the disease. However, it is important to recognize that FSHD is a complex disorder, and in addition to overexpression of 4q35 genes, factors such as the allelic variability of 4q35 genes, gender, and environment may also affect disease onset and severity.

The precise basis by which overexpression of 4q35 genes results in FSHD remains to be determined. Detailed analysis of the sequence surrounding the D4Z4 repeats has revealed it to be a relatively gene-poor region with a high density of pseudogenes and repetitive

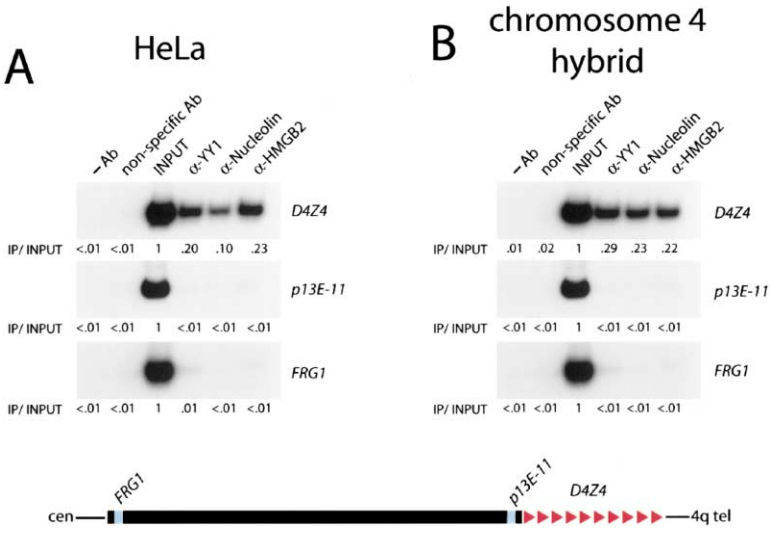

Figure 5. The DRC Binds D4Z4 In Vivo

(A) Chromatin immunoprecipitation (ChIP) analysis of D4Z4 binding in HeLa cells. Following formaldehyde crosslinking, purified chromatin was immunoprecipitated with antibodies to YY1, nucleolin, or HMGB2 or with a nonspecific antibody, and the purified DNA was analyzed by quantitative PCR using primers specific for D4Z4, p13E-11, or *FRG1*. The signal in the immunoprecipitate (IP) was quantitated by phosphoimager analysis and expressed as the fraction of the starting (INPUT) material.

(B) ChIP analysis in a human/rodent monochromosomal cell hybrid containing a single human chromosome 4 such that the only genomic copy of D4Z4 is present at human 4q35. Samples were treated and analyzed as described in (A). (Bottom) Schematic representation of the analyzed region showing the relative positions of D4Z4, p13E-11, and *FRG1*.

elements (van Geel et al., 1999; Blair et al., 2002). Although a number of short genes have been predicted by bioinformatic analyses, their existence has not been confirmed by either expressed sequence tag database searches or RT-PCR (our unpublished data).

Among the known 4q35 genes, *FRG1* and *FRG2*, which we have shown are overexpressed in FSHD muscle, lack homologies with other protein and contain no recognizable protein motifs. By far the best candidate for an FSHD-inducing gene is *ANT1*, which encodes an adenine nucleotide translocator. *ANT1* is a central component of the mitochondrial permeability transition pore complex and is predominantly expressed in heart and skeletal muscle. *ANT1* mutations have been found in patients with progressive external ophthalmoplegia, an autosomal dominant myopathy (Kaukonen et al., 2000; Napoli et al., 2001), and *ANT1* is upregulated in certain cardiomyopathies (Dorner and Schultheiss, 2000). Most importantly, *ANT1* overexpression induces apoptosis (Bauer et al., 1999); increased apoptosis is a feature of dystrophic muscles (Tews and Goebel, 1997), including FSHD muscle (Sandri et al., 2001). Muscle weakness, the cardinal feature of all myopathies, is due to reduced number of muscle fibers. Therefore, an attractive model for FSHD pathogenesis is that overexpression of *ANT1* induces apoptosis in muscle fibers, thereby reducing their number and resulting in disease.

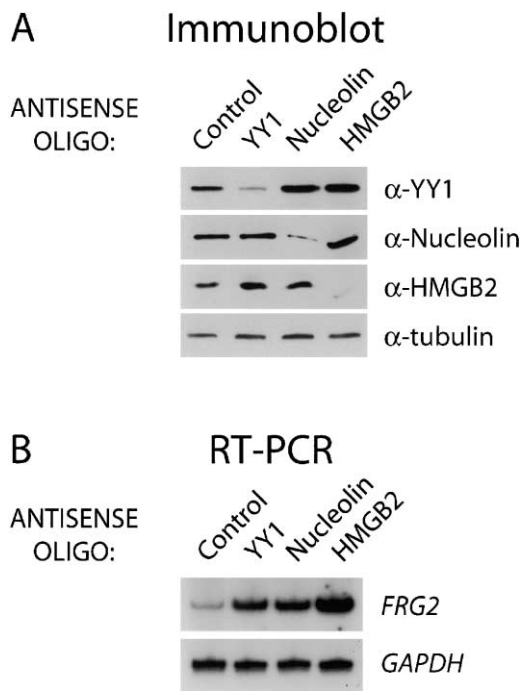

Figure 6. Reducing the Levels of DRC Components Results in Overexpression of 4q35 Genes

(A) HeLa cells were transfected with morpholino antisense oligonucleotides specific for YY1, HMGB2, or nucleolin or with a control morpholino oligonucleotide. Immunoblot analysis was performed to confirm the specific reduction in the level of each protein following transfection.

(B) *FRG2* and *GAPDH* expression analysis was performed by RT-PCR analysis on RNA extracted from the transfected cells.

The DRC Binds D4Z4 and Represses Transcription of 4q35 Genes

We found that the DRC is composed of YY1, HMGB2, and nucleolin and associates with D4Z4 sequences both in vitro and in vivo. In vitro binding assays demonstrated that association of the complex with D4Z4 was mediated through specific contacts between YY1 and a 27 bp element contained within D4Z4. These data suggest that YY1 is the major protein responsible for DNA binding, although other DRC components could contribute to DNA binding affinity and specificity. Significantly, decreasing the intracellular level of any one of the three DRC proteins derepressed 4q35 gene transcription, indicating that all three components are required for a functional silencing complex.

YY1, HMGB2, and nucleolin have been previously implicated in gene regulation and in particular transcriptional repression. YY1 is a complex protein involved in both repressing and activating a number of genes. The role of YY1 in repressing gene transcription is thought to be mediated through interactions with corepressors

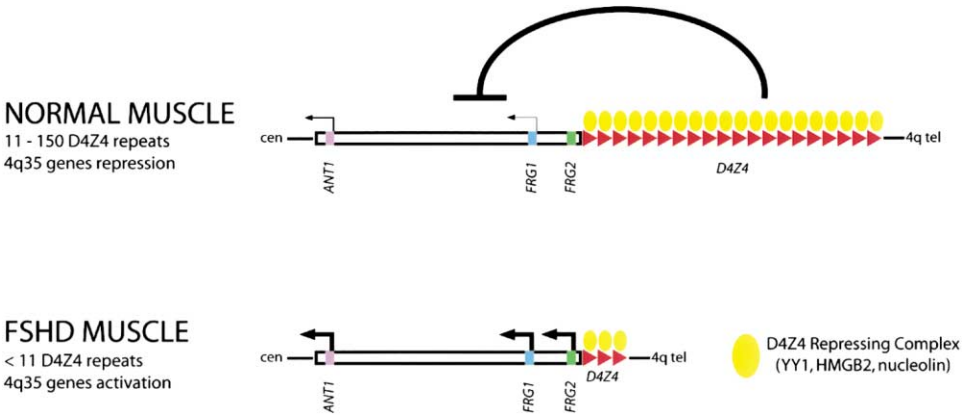

Figure 7. A Model for the Molecular Basis of FSHD

In normal individuals, the presence of a threshold number of D4Z4 repeats leads to repression of 4q35 genes by virtue of a DNA bound multiprotein complex that actively suppresses gene expression. In FSHD patients, deletion of an integral number of D4Z4 repeats reduces the number of bound repressing complexes and consequently decreases (or abolishes) transcriptional repression of 4q35 genes. As a result, these genes are inappropriately overexpressed, ultimately leading to disease onset and progression.

such as histone deacetylases (Thomas and Seto, 1999). HMGB2 is a member of a family of nonhistone nuclear proteins that modulate chromatin architecture (Thomas and Travers, 2001). Interestingly, HMGB2 has previously been shown to interact with two other proteins, SP100 and HP1, to form a transcriptional repressing complex (Lehming et al., 1998). Nucleolin has been shown to be a component of several transcription factor complexes (Hanakahi et al., 1997; Ying et al., 2000; Schulz et al., 2001) and to interact with other nuclear components potentially involved in the regulation of gene expression including YY1 (Xie et al., 1998).

The mechanistic basis by which the DRC represses transcription remains to be determined. One possibility is that the D4Z4 bound DRC could directly interact with and block the function of the general transcription machinery. Alternatively, the D4Z4 bound DRC could act by establishing a repressive chromatin structure. The fact that multiple 4q35 genes are repressed and that repression can occur over a long distance leads us to favor the possibility that D4Z4 initiates formation of a repressive chromatin structure that can be propagated to repress 4q35 genes located even at distal sites.

Parallels between Transcriptional Repression of 4q35 Genes and Yeast Telomeric Silencing

Transcriptional repression at D4Z4 bears several parallels to telomeric silencing in yeast. The most notable similarity is the presence of repetitive *cis*-acting DNA elements that function as binding sites for transcriptional repressing complexes. In yeast, the telomeric repeat sequences present at the chromosome terminus contain binding sites for the repressor/activator protein Rap1p. DNA bound Rap1p recruits additional factors, known as the Sir proteins, which in turn interact with specific histone isoforms to assemble a silencing complex (reviewed in Lustig, 1998). Similarly, each D4Z4 repeat unit contains a binding site for YY1, a multifunctional repressor/activator protein that associates with at least two other proteins, HMGB2 and nucleolin, to form a transcriptional repressing complex.

A second noteworthy similarity is that repression varies as a function of chromosome position. Yeast telomeric silencing is initiated at the chromosome terminus and, in general, decreases as a function of distance from the telomere (Renauld et al., 1993). Similarly, we have found a relationship between the distance of the gene from the 4q telomere and transcriptional repression (Figure 1). Recent studies have shown that silencing at yeast telomeres is discontinuous, with repression occurring only in limited domains (Pryde and Louis, 1999). Whether 4q35 genes in addition to *FRG1*, *FRG2*, and *ANT1* are repressed remains to be determined.

A Role for DNA Repetitive Elements

Our results provide insights into the biological function of DNA repetitive elements in gene transcription and their potential role in human diseases. D4Z4 has a VNTR structure: each 3.3 kb repeat contains two homeoboxes, and two different classes of human repetitive sequences (Hewitt et al., 1994). In FSHD, D4Z4 "short" alleles are strictly associated with the disease. Moreover, a group of VNTR alleles in the promoter of the insulin gene are found in linkage disequilibrium with Type 1 insulin-dependent diabetes mellitus (Bennett et al., 1995), correlated with insulin transcription level (Kennedy et al., 1995), and associated with birth size (Dunger et al., 1998). Similarly, a VNTR located downstream from the human proto-oncogene HRAS1 has been implicated in bladder cancer (Krontiris et al., 1993). It is tempting to speculate that these repetitive elements have a role in regulating gene expression, analogous to those in D4Z4.

Experimental Procedures

Cell Lines, Cell Culture

HeLa and C2C12 mouse myoblast cell lines were obtained from the American Type Culture Collection. GM10115 (Coriell Cell Repositories) is a human/chinese hamster somatic cell hybrid retaining human chromosome 4. HeLa and C2C12 cell lines were routinely cultured in a humidified atmosphere at 37°C with 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf

serum (FCS), 2 mM glutamine, and 70 μg/ml gentamycin. GM10115 cells were cultured in the same media containing 0.2 mM proline.

Antibodies

Primary antibodies used for immunoblotting and chromatin IP experiments were obtained as follows: α-YY1 and α-nucleolin from Santa Cruz, α-HMGB2 from BD Transduction Laboratories, and α-tubulin from Sigma. The 7G2 monoclonal α-nucleolin antibody was kindly provided by S. Pinol-Roma.

Purification of the D4Z4 Binding Proteins

HeLa nuclear extracts were prepared according to standard protocols (Ausubel et al., 2001) and dialyzed against dialysis buffer (20 mM HEPES [pH 8], 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT). Sixty milliliters of nuclear extract (360 mg of protein) was fractionated on 30 ml of P11 Phosphocellulose (Whatman) using a 0.1–1 M KCl step gradient in dialysis buffer. The active fraction was dialyzed against dialysis buffer and fractionated on 1 ml of DEAE-Sepharose (Pharmacia) using a 0.1–1 M KCl linear gradient in dialysis buffer. Active fractions were pooled, dialyzed against binding buffer 50 (20 mM HEPES [pH 8], 10% glycerol, 50 mM KCl, 0.2 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT), and loaded onto a DNA affinity resin prepared by coupling the D4Z4 minimal binding site to CNBr-activated Sepharose 4B (Pharmacia) according to standard protocols (Ausubel et al., 2001). This column was eluted using a 0.05–1 M KCl linear gradient in binding buffer 50. Active fractions were pooled, dialyzed against binding buffer 150 (20 mM HEPES [pH 8], 10% glycerol, 150 mM KCl, 0.2 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT), and fractionated on a new D4Z4 affinity column using a 0.15–1 M KCl linear gradient in binding buffer 150. For peptide sequencing, 0.1 ml of a final active fraction (representing ~1/20 of the final yield) was resolved by 12% SDS-PAGE. Following staining with Silver Staining Plus (BioRad), the protein bands were excised from the gel and analyzed by MALDI-TOF Mass Spectrometry at the University of Massachusetts Medical School Protein Microsequencing and Mass Spectrometry Center.

Immunoblot Analysis

Total lysates were obtained by boiling the samples in 1× SDS sample buffer (Laemmli buffer), fractionated onto SDS-PAGE, and electroblotted to PVDF filters (Millipore) in 25 mM Tris base, 192 mM glycine, 20% v/v methanol (pH 8.3) using a TransBlot cell (BioRad). The membrane was blocked in 5% nonfat dry milk (Carnation) in TBST (20 mM Tris-HCl [pH 7.5], 150 mM NaCl, 0.1% Tween-20) for 60 min, followed by incubation with the primary antibody for 60 min. After three serial washings with TBST for 10 min, the antigen-antibody complexes were visualized with the appropriate secondary antibody (Jackson ImmunoResearch) conjugated to horseradish peroxydase and a chemiluminescent system as recommended by the manufacturer (NEN).

Electro-Mobility Shift Assay (EMSA) and Supershift

The KpnI-D4Z4 unit plasmid, used to generate EMSA probes for the experiments shown in Figure 2, was kindly provided by L. Felicetti. For the EMSA experiments shown in Figure 4, oligonucleotide probes were prepared by end-labeling double-stranded oligonucleotides covering the region of the minimal binding site. The sequences of the oligonucleotides were as follows: WT, 5′-CTCACCG CCATTCATGAAGGGGTGGAGCCTGCCTG-3′; MUT, 5′-CTCACCG CC*gcc*CATGAAGGGGTGGAGCCTGCCTG-3′ (mutated nucleotides are reported in italics). DNA binding reactions were carried out in 20 μl of 10 mM HEPES (pH 8), 10% glycerol, 50 mM KCl, 0.1 mM EDTA, 0.25 mM DTT, 0.1 mg/ml poly(dI-dC), 0.1 mg/ml bovine serum albumin, 20,000 cpm of radiolabeled probe. The samples were incubated 30 min on ice in the presence of the indicated nuclear extract or protein fraction. For competition assays, unlabeled probes were added to the reaction mixture at the same time as the labeled probe. For supershift assays, binding reaction was assembled in the absence of radiolabeled probe. Antibodies were added to the binding reaction and the samples were incubated 15 min at room temperature. Radiolabeled probe was added and the samples were incubated for additional 30 min on ice. Reaction mixtures were loaded in the cold room onto 6% native polyacrylamide gels containing 0.5× Tris-borate-EDTA (TBE) that had been preelectrophoresed for 30 min at 20 mA. After electrophoresis for 1.5–2 hr at 20 mA, the gels were dried and exposed for autoradiography.

DNase I In Vitro Footprinting

DNase I in vitro footprinting experiments were performed essentially as described (Ausubel et al., 2001). D4Z4-243 cloned in pBluescript SK^+ vector (Stratagene) was used as probe. Briefly, the vector was digested with Asp718, labeled with Klenow in the presence of [α-^{32}P]dATP and [α-^{32}P]dCTP, digested with SacI, and gel purified. DNA binding reactions were carried out as for EMSA in the presence of different amounts of nuclear extracts. After a 30 min incubation on ice, 0.3 units of DNase I (RQ1, Promega) was added and the samples incubated 1 min at room temperature. Reactions were stopped, and phenol was extracted, precipitated, and resuspended in formamide loading buffer. Samples were loaded onto 6% denaturing polyacrylamide gels along with sequencing reactions prepared as described (Ausubel et al., 2001). After electrophoresis the gels were dried and exposed for autoradiography.

Chromatin Immunoprecipitation (ChIP)

ChIP analysis was performed according to published procedures (Boyd et al., 1998; the protocol was kindly provided by P.J. Farnham). Quantitative PCR was performed as previously described (Orlando and Paro, 1993) using an aliquot (1/30) of the purified DNA. PCR reactions contained [α-^{32}P]dATP (2.5 μCi for each 25 μl reaction), and the PCR products were detected by autoradiography after separation on a 6% polyacrylamide gel. Sequences of the oligonucleotides used are as follows: D4Z4 region, DBS-I (5′-AGGCCTCGACGCCCTGG GTC-3′) and DBS-II (5′-TCAGCCGGACTGTGCACTGCGGC-3′); p13E-11 region, p13-I (5′-AGCCCTGCCACAGGCTTCTGTG-3′) and p13-II (5′-AGTGCTTATGCCTGAGGAATCTG-3′); FRG1, FRG1-1f (5′-TCTACAGAGACGTAGGCTGTCA-3′) and FRG1-1rb (5′-CTTGAG CACGAGCTTGGTAG-3′).

RNA Extraction and RT-PCR

Total RNA was prepared using the TRI Reagent (Sigma) according to the manufacturer's instructions for the isolation of RNA for RT-PCR. Purified RNA was treated with RNase-free DNase I (Promega) to remove residual DNA, and 1 μg of purified DNA-free RNA was used for first-strand cDNA synthesis with SuperScript II RNase H^- RT (Invitrogen) according to the manufacturer's instructions. Quantitative PCR was performed as previously described (Orlando and Paro, 1993) using an aliquot (1/20) of the RT reaction. PCR reactions contained [α-^{32}P]dATP (2.5 μCi for each 25 μl reaction), and the PCR products were detected by autoradiography after separation on a 6% polyacrylamide gel. Sequences of the oligonucleotides used are as follows: FRG1, FRG1-1f (5′-TCTACAGAGACGTAGGCTGTCA-3′) and FRG1-1rb (5′-CTTGAGCACGAGCTTGGTAG-3′); FRG2, EX2BF (5′-CCAGAGTCCAGCTCATATCG-3′) and SSC8 (5′-CTCACAGGTA AGTGGAGAATGG-3′); ANT1, ANT1 (5′-GTGCATTAAGTGGTCTTT ATT-3′) and ANT2 (5′-TGTGGTTTAATAGACTATTCCTA-3′); GAPDH, G3PDH 5′ (5′-ACCACAGTCCATGCCATCAC-3′) and G3PDH 3′ (5′-TCCACCACCCTGTTGCTGTA-3′). The FRG2 oligos were kindly provided by R.R. Frants and S. Van der Maarel. FRG1 and ANT1 primers are 4q35 specific. Closely related sequences to FRG2 are present on chromosomes 1, 4, 8, 10, and 20; however, the FRG2 copy at 4q35 is the only one resistant to digestion with TaiI and HhaI (R.R. Frants, personal communication). Therefore, to obtain a 4q35-specific signal, FRG2 PCR products were digested with both enzymes before electrophoresis.

Purification of GST-YY1

GST-YY1 expression vectors were kindly provided by Y. Shi, E. Seto, A. Usheva, D.M. Margolis, and T. Shenk. For purification, saturated cultures of *E. coli* expressing the GST-YY1 fusion protein were diluted 1:20 with fresh media and incubated at 37°C until OD at 600 nm was 0.6. IPTG was added to a final concentration of 0.2 mM, and the bacteria were incubated at 37°C for 4 hr. Bacteria were harvested, washed once with STE (10 mM Tris [pH 8.0], 150 mM NaCl, 1 mM EDTA), and resuspended in STE containing 100 μg/ml of lysozyme. Following a 15 min incubation on ice, DTT was added to a final concentration of 5 mM. N-Lauryl Sarcosine was added to

a final concentration of 1.5% (from a 10% stock in STE), and the bacteria were sonicated on ice until the solution was clear (~1 min). Triton X-100 was added to a final concentration 3%. Following centrifugation, the supernatant was added to glutathione agarose beads (50% suspension in STE) for affinity purification. Fusion proteins were eluted using 75 mM HEPES (pH 7.4), 150 mM NaCl, 10 mM reduced glutathione, 5 mM DTT, and 2% N-octyl glucoside. Purified GST-YY1 was dialyzed against dialysis buffer (20 mM HEPES [pH 8], 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT).

Antisense Morpholino Oligonucleotides

Morpholino oligonucleotides (GeneTools) spanning the sequence of the AUG translational start site of human YY1 (5′-CATGGCTGAGGG CTCCGCCGCCACG-3′), HMGB2 (5′-GGGTCTCCTTTACCCATGTT GACAG-3′), nucleolin (5′-GCGAGCTTCACCATGATGGCGGCGG-3′), or standard control oligonucleotides (5′-CCTCTTACCTCAGTTA CAATTTATA-3′) were introduced into HeLa cells according to the manufacturer's special delivery protocol (Morcos, 2001). Oligonucleotide delivery was repeated at 48 hr intervals for a total of five times. Cells were maintained at exponential growth throughout the entire procedure. Immunoblotting and quantitative RT-PCR analysis was performed 48 hr after the last delivery.

Repression Assay

The vector pcDNA 3.1 (Invitrogen) was digested with KpnI and SmaI, blunt-ended, and religated to obtain CMV-Neo. Tandem repeats of the DBE sequence were produced as described (Ausubel et al., 2001) and cloned into the NheI site of CMV-Neo to obtain CMV-DBE-Neo. A spacer sequence equivalent in length to six DBE repeats was generated by digesting pBluescript SK+ (Stratagene) with SpeI and PvuII and was cloned into NheI/HindIII-digested CMV-Neo that had been blunt-ended at the HindIII site to obtain CMV-Spacer-Neo. Each construct was verified by sequencing.

For the repression assay, each construct was linearized with PvuI and transfected into HeLa cells in 6-well plates with Effectene according to the manufacturer's instructions (Qiagen). Forty-eight hours after transfection, one-third of each well was passed in 100 mm plates, and G418 (1 mg/ml, Calbiochem) was added to the media. Two weeks later, colonies were fixed and stained with Giemsa, and the number of G418-resistant cells counted. Experiments were repeated in triplicate, using two different plasmid preparations.

Acknowledgments

We are indebted to all FSHD families and the FSH Society. We thank S. Bacchetti, E. Green, and G. Perini for their in-depth critique of the manuscript; C. Ghigna, P. Norio, and past and present members of the M. Green laboratory for helpful discussions; and S. Evans for editorial assistance. We thank C. Angelini (Telethon Tissue Bank C7) for providing muscle tissues and C. Danesino (Telethon Cell Bank) for providing peripheral blood lymphocytes. We also gratefully acknowledge R.R. Frants for permission to describe unpublished results. We apologize in advance to those whose work we were not able to cover due to space constraints. This work was supported by a Marjorie Bronfman-FSH Society fellowship to D.G. and the Worcester Foundation Award and Legato Ferrari Award as well as grants from the National Institutes of Health (R21 NS43973) and the Muscular Dystrophy Association and Telethon (GP284Y01) to R.T. M.R.G. is an investigator of the Howard Hughes Medical Institute.

Received: March 4, 2002
Revised: June 17, 2002

References

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (2001). Current Protocols in Molecular Biology (New York: John Wiley & Sons).

Bauer, M.K., Schubert, A., Rocks, O., and Grimm, S. (1999). Adenine nucleotide translocase-1, a component of the permeability transition pore, can dominantly induce apoptosis. J. Cell Biol. *27*, 1493–1502.

Bennett, S.T., Lucassen, A.M., Gough, S.C., Powell, E.E., Undlien, D.E., Pritchard, L.E., Merriman, M.E., Kawaguchi, Y., Dronsfield, M.J., Pociot, F., et al. (1995). Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. Nat. Genet. *9*, 284–292.

Blair, I.P., Adams, L.J., Badenhop, R.F., Moses, M.J., Scimone, A., Morris, J.A., Ma, L., Austin, C.P., Donald, J.A., Mitchell, P.B., and Schofield, P.R. (2002). A transcript map encompassing a susceptibility locus for bipolar affective disorder on chromosome 4q35. Mol. Psychiatry, in press.

Boyd, K.E., Wells, J., Gutman, J., Bartley, S.M., and Farnham, P.J. (1998). c-Myc target gene specificity is determined by a post-DNA binding mechanism. Proc. Natl. Acad. Sci. USA *95*, 13887–13892.

Chung, J.H., Bell, A.C., and Felsenfeld, G. (1997). Characterization of the chicken β-globin insulator. Proc. Natl. Acad. Sci. USA *94*, 575–580.

Clark, L.N., Koehler, U., Ward, D.C., Wienberg, J., and Hewitt, J.E. (1996). Analysis of the organisation and localisation of the FSHD-associated tandem array in primates: implications for the origin and evolution of the 3.3 kb repeat family. Chromosoma *105*, 180–189.

Dorner, A., and Schultheiss, H.P. (2000). The myocardial expression of the adenine nucleotide translocator isoforms is specifically altered in dilated cardiomyopathy. Herz *25*, 176–180.

Dunger, D.B., Ong, K.K., Huxtable, S.J., Sherriff, A., Woods, K.A., Ahmed, M.L., Golding, J., Pembrey, M.E., Ring, S., Bennett, S.T., and Todd, J.A. (1998). Association of the INS VNTR with size at birth. ALSPAC study team. Avon longitudinal study of pregnancy and childhood. Nat. Genet. *19*, 98–100.

Eng, J.K., McCormack, A.L., and Yates, J.R., Iii. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass Spectrom. *5*, 976–989.

Fisher, J., and Upadhyaya, M. (1997). Molecular genetics of facioscapulohumeral muscular dystrophy (FSHD). Neuromuscul. Disord. *7*, 55–62.

Gabriels, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D., and Belayew, A. (1999). Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. Gene *236*, 25–32.

Goto, K., Lee, J.H., Matsuda, C., Hirabayashi, K., Kojo, T., Nakamura, A., Mitsunaga, Y., Furukawa, T., Sahashi, K., and Arahata, K. (1995). DNA rearrangements in Japanese facioscapulohumeral muscular dystrophy patients: clinical correlations. Neuromuscul. Disord. *5*, 201–208.

Hanakahi, L.A., Dempsey, L.A., Li, M.J., and Maizels, N. (1997). Nucleolin is one component of the B cell-specific transcription factor and switch region binding protein, LR1. Proc. Natl. Acad. Sci. USA *94*, 3605–3610.

Hewitt, J.E., Lyle, R., Clark, L.N., Valleley, E.M., Wright, T.J., Wijmenga, C., van Deutekom, J.C., Francis, F., Sharpe, P.T., Hofker, M., et al. (1994). Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. Hum. Mol. Genet. *3*, 1287–1295.

Hsu, Y.D., Kao, M.C., Shyu, W.C., Lin, J.C., Huang, N.E., Sun, H.F., Yang, K.D., and Tsao, W.L. (1997). Application of chromosome 4q35-qter marker (pFR-1) for DNA rearrangement of facioscapulohumeral muscular dystrophy patients in Taiwan. J. Neurol. Sci. *149*, 73–79.

Kaukonen, J., Juselius, J.K., Tiranti, V., Kyttala, A., Zeviani, M., Comi, G.P., Keranen, S., Peltonen, L., and Suomalainen, A. (2000). Role of adenine nucleotide translocator 1 in mtDNA maintenance. Science *289*, 782–785.

Kennedy, G.C., German, M.S., and Rutter, W.J. (1995). The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. Nat. Genet. *9*, 293–298.

Krontiris, T.G., Devlin, B., Karp, D.D., Robert, N.J., and Risch, N. (1993). An association between the risk of cancer and mutations in the HRAS1 minisatellite locus. N. Engl. J. Med. *329*, 517–523.

Lehming, N., Le Saux, A., Schuller, J., and Ptashne, M. (1998). Chro-

matin components as part of a putative transcriptional repressing complex. Proc. Natl. Acad. Sci. USA *95*, 7322–7326.

Lunt, P.W. (1998). 44th ENMC international workshop: facioscapulohumeral muscular dystrophy: molecular studies 19–21 July 1996, Naarden, The Netherlands. Neuromuscul. Disord. *8*, 126–130.

Lunt, P.W., Jardine, P.E., Koch, M.C., Maynard, J., Osborn, M., Williams, M., Harper, P.S., and Upadhyaya, M. (1995). Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35-facioscapulohumeral muscular dystrophy (FSHD). Hum. Mol. Genet. *4*, 951–958.

Lustig, A.J. (1998). Mechanisms of silencing in *Saccharomyces cerevisiae*. Curr. Opin. Genet. Dev. *8*, 233–239.

Lyle, R., Wright, T.J., Clark, L.N., and Hewitt, J.E. (1995). The FSHD-associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subsets of which are clustered on the short arms of the acrocentric chromosomes. Genomics *28*, 389–397.

Morcos, P.A. (2001). Achieving efficient delivery of morpholino oligos in cultured cells. Genesis *30*, 94–102.

Napoli, L., Bordoni, A., Zeviani, M., Hadjigeorgiou, G.M., Sciacco, M., Tiranti, V., Terentiou, A., Moggio, M., Papadimitriou, A., Scarlato, G., and Comi, G.P. (2001). A novel missense adenine nucleotide translocator-1 gene mutation in a Greek adPEO family. Neurology *57*, 2295–2308.

Orlando, V., and Paro, R. (1993). Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin. Cell *75*, 1187–1198.

Padberg, G.W. (1982). Facioscapulohumeral disease. MD Thesis, Leiden University, Leiden, The Netherlands.

Pryde, F.E., and Louis, E.J. (1999). Limitations of silencing at native yeast telomeres. EMBO J. *18*, 2538–2550.

Renauld, H., Aparicio, O.M., Zierath, P.D., Billington, B.L., Chhablani, S.K., and Gottschling, D.E. (1993). Silent domains are assembled continuously from the telomere and are defined by promoter distance and by SIR3 dosage. Genes Dev. *7*, 1133–1145.

Ricci, E., Galluzzi, G., Deidda, G., Cacurri, S., Colantoni, L., Merico, B., Piazzo, N., Servidei, S., Vigneti, E., Pasceri, V., et al. (1999). Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. Ann. Neurol. *45*, 751–757.

Sandri, M., El Meslemani, A.H., Sandri, C., Schjerling, P., Vissing, K., Andersen, J.L., Rossini, K., Carraro, U., and Angelini, C. (2001). Caspase 3 expression correlates with skeletal muscle apoptosis in Duchenne and facioscapulo human muscular dystrophy. A potential target for pharmacological treatment? J. Neuropathol. Exp. Neurol. *60*, 302–312.

Sarfarazi, M., Wijmenga, C., Upadhyaya, M., Weiffenbach, B., Hyser, C., Mathews, K., Murray, J., Gilbert, J., Pericak-Vance, M., Lunt, P., et al. (1992). Regional mapping of facioscapulohumeral muscular dystrophy gene on 4q35: combined analysis of an international consortium. Am. J. Hum. Genet. *51*, 396–403.

Schulz, M., Schneider, S., Lottspeich, F., Renkawitz, R., and Eggert, M. (2001). Identification of nucleolin as a glucocorticoid receptor interacting protein. Biochem. Biophys. Res. Commun. *280*, 476–480.

Tawil, R., Forrester, J., Griggs, R.C., Mendell, J., Kissel, J., McDermott, M., King, W., Weiffenbach, B., and Figlewicz, D. (1996). Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY Group. Ann. Neurol. *39*, 744–748.

Tews, D.S., and Goebel, H.H. (1997). DNA-fragmentation and expression of apoptosis-related proteins in muscular dystrophies. Neuropathol. Appl. Neurobiol. *23*, 331–338.

Thomas, M.J., and Seto, E. (1999). Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? Gene *236*, 197–208.

Thomas, J.O., and Travers, A.A. (2001). HMG1 and *2*, and related 'architectural' DNA-binding proteins. Trends Biochem. Sci. *3*, 167–174.

Tupler, R., Berardinelli, A., Barbierato, L., Frants, R., Hewitt, J.E., Lanzi, G., Maraschio, P., and Tiepolo, L. (1996). Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. J. Med. Genet. *33*, 366–370.

Tupler, R., Perini, G., Pellegrino, M.A., and Green, M.R. (1999). Profound misregulation of muscle-specific gene expression in facioscapulohumeral muscular dystrophy. Proc. Natl. Acad. Sci. USA *96*, 12650–12654.

van Deutekom, J.C., Lemmers, R.J., Grewal, P.K., van Geel, M., Romberg, S., Dauwerse, H.G., Wright, T.J., Padberg, G.W., Hofker, M.H., Hewitt, J.E., and Frants, R.R. (1996). Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35. Hum. Mol. Genet. *5*, 581–590.

van Geel, M., Heather, L.J., Lyle, R., Hewitt, J.E., Frants, R.R., and de Jong, P.J. (1999). The FSHD region on human chromosome 4q35 contains potential coding regions among pseudogenes and a high density of repeat elements. Genomics *61*, 55–65.

van Geel, M., van Deutekom, J.C., van Staalduinen, A., Lemmers, R.J., Dickson, M.C., Hofker, M.H., Padberg, G.W., Hewitt, J.E., de Jong, P.J., and Frants, R.R. (2000). Identification of a novel beta-tubulin subfamily with one member (TUBB4Q) located near the telomere of chromosome region 4q35. Cytogenet. Cell Genet. *88*, 316–321.

Wijmenga, C., Hewitt, J.E., Sandkuijl, L.A., Clark, L.N., Wright, T.J., Dauwerse, H.G., Gruter, A.M., Hofker, M.H., Moerer, P., Williamson, R., et al. (1992). Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. Nat. Genet. *2*, 26–30.

Winokur, S.T., Bengtsson, U., Feddersen, J., Mathews, K.D., Weiffenbach, B., Bailey, H., Markovich, R.P., Murray, J.C., Wasmuth, J.J., Altherr, M.R., et al. (1994). The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. Chromosome Res. *2*, 225–234.

Xie, J., Briggs, J.A., and Briggs, R.C. (1998). Human hematopoietic cell specific nuclear protein MNDA interacts with the multifunctional transcription factor YY1 and stimulates YY1 DNA binding. J. Cell. Biochem. *70*, 489–506.

Yant, S.R., Zhu, W., Millinoff, D., Slightom, J.L., Goodman, M., and Gumucio, D.L. (1995). High affinity YY1 binding motifs: identification of two core types (ACAT and CCAT) and distribution of potential binding sites within the human beta globin cluster. Nucleic Acids Res. *23*, 4353–4362.

Ying, G.G., Proost, P., van Damme, J., Bruschi, M., Introna, M., and Golay, J. (2000). Nucleolin, a novel partner for the Myb transcription factor family that regulates their activity. J. Biol. Chem. *275*, 4152–4158.

Zatz, M., Marie, S.K., Passos-Bueno, M.R., Vainzof, M., Campiotto, S., Cerqueira, A., Wijmenga, C., Padberg, G., and Frants, R. (1995). High proportion of new mutations and possible anticipation in Brazilian facioscapulohumeral muscular dystrophy families. Am. J. Hum. Genet. *56*, 99–105.